

FIG. 1A

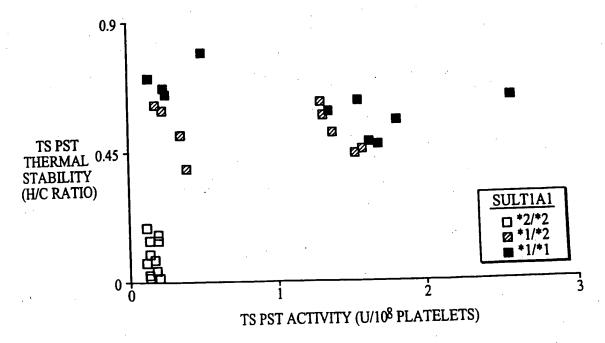
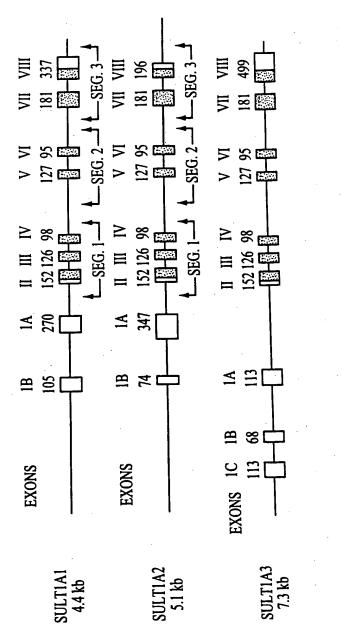
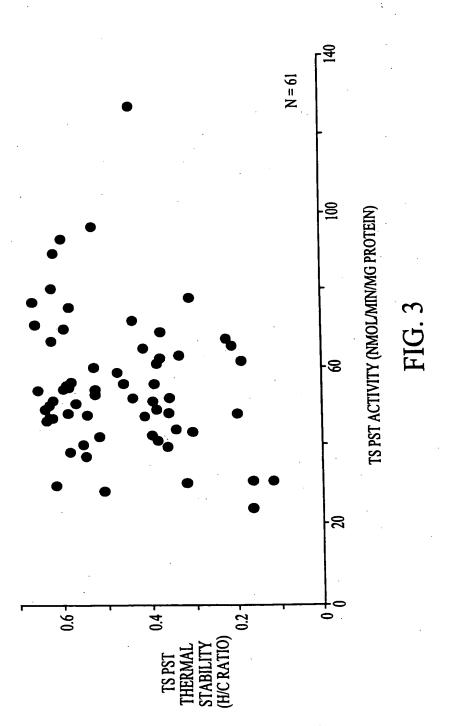


FIG. 1B





Applicant(s): Richard M. Weinshilboum et al. SULFOTRANSFERASE SEQUENCE VARIANTS

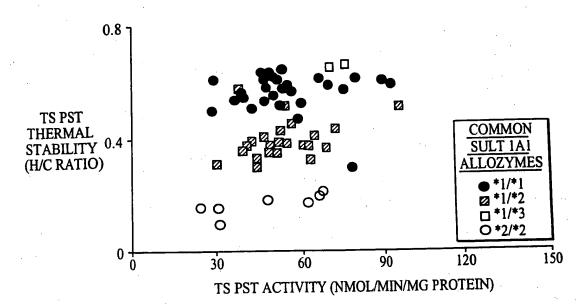


FIG. 4A

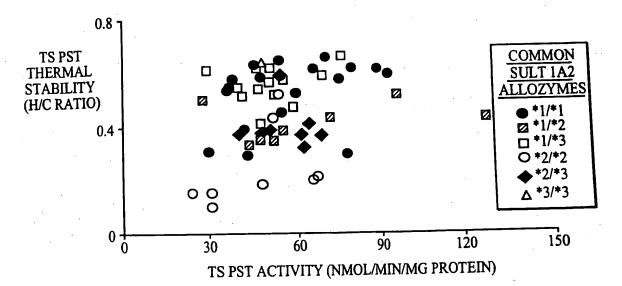


FIG. 4B

ccaaccctgg

tggagatoso acconsoco ttoottoogg aagoaaatot aagtocagoo pototogtot aggtcoccag atgoccetto \*\*\*\*\*\*\*\*\*\*\*\*\* atgaatgtag tgacotttat agegeestgt gtccgacagg ccagtcacgg ottococcao gtgoacagot gasococtgg gectoggitos ccccaaagtg satscaggtg tetteteete aggottgact aggtoatcae asaccggtgg cagtococcg agcctggagg toatggatga ctgtgaccaa aatggogtga gtogtggott tgetecacoo gotoggasas sasassasas cccattgtgg catgootoag tacotgtoto tggottocaa acotoggoot tottotates atgittototg tgcasaagtc ggtgtggggc ggtgccacaa tttacaccca catcasatec csagettect tgaggacctg csotdggccs tocastocct cocctotett cctgggotgc getggagtge tgootogged tagctgagtg ttosiggetg gtgtgcoacc gatetggetg agossactes ctgtgggotg stocstgoto gtgatotgoo taaggacaag gtgasscots geagetattt getagtagee otggsattco gatggttgtt tgoootsoss tgttgcccag gattecagge ttetgtgtes gocastgoco gatottggta tgggtgacot cactocotgt ggcotagggg ttttoteoec \*gotttotoo gegtgtotee gagatectag otogoctett cotttactos tgttotatgo totaggggo cocadtgaca agasscaatt caggggtgga totoctotgo acgggoootg teasgacegt tottotagge agatactggt gtttosotot aagtagetgg cogsettesg cttagcgcat acadadacco geogtateet gtoactgett ctccctacco occetagagg osttottotg tatottacid atcatacast gtgcatgggo tgcotatoco aggtoaggag 0000000000 tasagattto cagetiggie etestectat titteaggga ettatiggee stotoscatt atgaga agod gggagattos poppopted gggttgtggt tgttctttct teagectece ttgotgaggg asteceetee totosssoto caccttagec gtoctcaaga getgtgagge tggccccagc ttgagaccga geagettotg фдесфадася gacaggatga ttockgggcd tttococttt gagaagotoo gttcaggggo otoctgaggg gccotagoat tggaagtott tagggggatg cacctgooad geagtgteet agagootgga caggaaaagg gttcaagcga ttotcctgco ctcostgttg gtosggotgg ggeettetee aggetettgg gascgtgggg cacttgctac cactotatta cagtgtocct tttttgttt tttttttt gtgcaatcca tgocttggta gagtgggett totoctog вадавосадд ccattgggcc goscactasg gagggtsatg cottgtggto taggtgacct tacggtttca tcaagggaac acgeatttot **க**்ஷே ஏருக ஏருக gasatasaso caatgggact atgasgagto ggggtocatg deddiffedda tossgsagag otcottgtot ettacotgco gtgcccagtt tetggeecea tgctggagct gtgagaggag ttcottccgg tagaagaogt cagasasas sacctacasa asosasocca gggtatatto otggtgggtg attttgtatt asagtettag ಕ್ಷರ್ಭಕ್ಷಗಳಿಗೆ ಪ್ರಕೃತ್ತಿಗಳಿಗೆ ototgaaoto aggateaga totagactgg agccagcagg tggotgootg aggaotese Caacccagg actoccaggg gactoggaat tggatagece tocaogettt paccatacca tgggaagagg tgtetecete ogcotcotgg gacaaggttt totocagotg gettgtetgt sctggcctgg tgcagagtaa agtgcaatac gtocgottgo ctttaaagtg gagootoott getggggetg Cacacatooc taggeses gs tgcotgaoto teagecegtt cogosacoto ogtettgges tggtcgtggc getgeotggg aatagtggaa aggagtgago ggtccagctt geaggeacae ttttagtaga sactasagsg tgatctgcot ctgoctotoc attagagaga tttaagottg caccttoaag ottgtotgtg actetgecea tattactag radccatgas ctgggattac ggettette gacategace tettggetea aattttotat geagatttoo ttoctggage gtgtctggt totossot gattoctca acteceacet conocttggs agageceaeg cotegettet pococagoot gggettgtga gcscgggggs ttgctgccag 2101 2201 2301 2401 2001 1301 1901 1001 101 1201 1401 501 100 1801 901 601 501 701 801 401 601

## FIG. 5B

	5	<b>-</b>		Ę		Z		
EACIN	Z	EVC		EXON		EXON		
THETTECCCE  TAR	GAAGTTCATG K F H tgcotgotto	CATGAAGGAB H K E gcagccccoa	aagaagacca gggggttgag	teegageete GGACTTCATG D F H	cccrcarGa PFHR atgagactoc	TATGCGGAGA Y A E K CAAGCGGCTC	ccassastr ggcoaggeat	FIG. 5C
oagereercr a	GCTTCCTGGA F.L.E tggoaccoct	rcrarcaaca y E D coctcocttg	ggototete acagteatgg otggggooot	gasteagess AGGAGACCGT E T V	CASCATCTCC S I S otocootggg	CGATGCGGAC D A D CAAACCTGAC	rcarrreage tatgattgtg	E E
cotdoccotd	ACCTGGGACA I W D S	CTCTACCTCT L Y L E Coaccagec		rccreccas s 1 P E	TCATGGACCA H D H cactcggggo	ATGAGCGCTT E R F GTGTGCGAAT	GGTGGGAGGA. gagaataaaa	
gttgegtot (	ccactrotac cacargecca aberecacc reascorese accressaca ectromessa saartcare H F Y H H A K Y H P E P O T W D S F L E K F H gaaagtgreeg a scootgree type coot type type type typesce a segmanda saccetyrist typesceet type typesceet	CGTGCAGGAG TGGTGGGAGC TGAGCCGCAC CCACCCTGTT CTCTACCTCT TCTATGAAGA CATGAAGGAD:  Y Q E W W E L S R T H P V L Y L Y E D H K E  Y Q E W W E L S R T G R T CACCCGGG COCCCGG COCCCGG CCCCGGGGGGGGGG	tignocady octoocgot acadgococa cotggococacaggoco acadteatog asgasgaca asacttagto etattgotae tocctgocaa agggiggoco accagggoco acagteatog asgasgaca asacttagto etattgotae	agotoatgat gggatoacaag tlagagaagg ggaccottt tactttoct gaatcageaa tecgageete EXON VII tggccagate geetetgagg tlagagaagg ggaccottt tactttoct gaatcagaa tecgageete EXON VII caalageeae Alfcaalaga Icclesaer refedeecac Icclescae Alfcaalaga Icclesaer refedeecac Icclescae Alfcaalaga Alfcaalaga Icclesaer refedeecac Icclescae Alfcaalaga Alfcaalaga Icclescaer refedeecac Icclescaer Icc	K R E I U K I L CACCACCOTC CCCCAGGAGT TCATGGACCA CAGCATCTCC CCCTTCATGA AAGAACCCTA TGACCAACTA CACCACGATC CCCCAGGAGT TCATGGACCA CAGCATCTCC CCCTTCATGA AAGAACCCTA TGACCAACGACCTC CCCCAGGAGT TCATGGACCA CAGCATCTCC CCCTTCATGA AAGAACCCTA TGACCAACGAC CCCCAGGAGT TCATGGACCA CAGCAACGA CAGAGACCA CAGAGACA CAGACA CAG	ATGGCTGGGG ACTGGAAGAC CACCTTCACC GTGGCGCAGA ATGAGCGCTT CGATGCGGAC TATGCGGAGA EXON VIII  ATGGCTGGGG ACTGGAAGAC CACCTTCACC GTGGCGCAGA ATGAGCGCTT CGATGCGGAC TATGCGGAGA EXCTT T T T T A Q N E R F D A D Y A E K  H A Ø D W K T T T T T T T T T T T T T T T T T T	TGGAGGCTGA	
atggtgotgg (	AGGTGCACCC V K P tggtggctac	TEASCCGCAC S R T C C C C C C C C C C C C C C C C C C	tocetgeesa	ttagagaagg TCTGGAGTT	CACCACGTC T T T	CACCTTCACC T F T	CCCABCAATT ATAAAACCAA	
stgtgtgtg (	CACATGGCCA H K A K agcogagggg	TGGTGGGAGC	ctattoctac ctattgctac	gggatoacag gcctctgagg ATTCAAAAGA	TGACCAACTA T N Y	ACTGGAAGAC	ASAIGGGGCTC TGTCTGTAAT ATAAAATAAA	
otggggoagg (	CCACTTCTAC H F Y	CGTBCAGGAG  V Q E	ttggmocagg seacttagto	ageteatgat tggccagate CAAAAGGGAG	K R E AAGAACCTA K N P H	ttggggoggg ATGGCTGGGG H A 0 D	CTGAGCTGTG E L CGGTAGGTCA.	
tgagetgge	FTTCCTACTA S Y Y Gggaggaag	SGTACCAGCA Y Q K		agccaagoco otaggagagg otoagaacc	N P GGAGATGAAG E H K	occccaggc	ABCTTCCGCT S F R S BGGCCTGGGA	ALI TOTAL TO
atocaagag (	ATGTGGCAG T A W rigggtttga	E T FIGURE GETACCASCA ST T Q R	stgtgaagot gattcccca satqqqqqc	secostagg geagggtet	COTCGTTCAA S F K	jtgotggcoa cotgcottoo	CTGCAGCCTC C S L TATGAATTGA	<u>recectaerr</u> g 7152
HANI offections astocaagag atgagetgge etgagggeagg etgagggggggggg	5601 CAACGCAAAG GATGTGGCAG TTTCCTACTA N A N D V A V S Y Y 5701 GTCGGAGAAG GTGGGGTttga tgggaggagg	toologics to the state of the second of the state of the	gtgagaocae etgtgaaget teettuary cageaggee ggatteeca teetgeette marksmitg categggage caeggateag	teacggtoot caccoatagg agceaageco gagttggote tgeagggtet etaggagg	GTTCAGCACA COTCOTTCAA GGAGATGAAG AAGAACCCTA TGACCACCGTC CCCCAGGAGT TCATGGACCA CAGCATCTCC CCCTTCATGA  GTTCAGCACA COTCOTTCAA GGAGATGAAG AAGAACCCTA TGACCACCACCTC CCCCAGGAGT TCATGGACCA CAGCATCTCC CCCTTCATGA  GTTCAGCACA COTCOTTCAA GGAGATGAAG AAGAACTA TATT T T T T T T T T T T T T T T T	GGAAAGHtgg gtgotggoom gtaogggggt	6801 AGATGGCAGG CTGCAGCCTC AGCTTCCGCT CTGAGCTGTG ASAGGGCTC CLGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	estagotost g 7152
10.52	5601		5901.0		6501	6501	6801.	7101

gggggaaata tegaceteag tggagccgtc gassctgotg cagoctggtc gtttetgeag cactagagtg tcotcagggg ctotgagete atttectges tgaggacaaa ttotatttt ttcttcaact totggttggt ttcapaccca ttoattotto tgttttotac gattacagg tesetecaga ccccaccc caaccotacc ttgtgatgo gasagettgt ettttggagt occacdtga tggctaagag cocceddocoo ggatgageac gaccascctt cagtgggaca ttgactaatt asagtgetgg acgtaggeag otttatggge caggtgttco ocatgttgot ctttaaatag tggageetee gogtgatett ctggtcatgt ggotggtttt toetggggat occacaodca coccapttgo gaegotgggg caagactggo aggetgagee tggctgtgat catcaaagao gccaccaggo toggoctoco ctacaaggtg tggggetgao gacotdccag ggctgcctgt gagtgcaatg aaagtcatga gacogiggity ocacaattta cacccaaata toctaggtga asoctacate asataoageg tttccttctt ctgotggtcc cctcagtttt tgaggtcccc agggagccat cttcaototg ctggcacctt ccctggctgt tgagtgcagt tgaccetgga caagcotosa cotoctotgg tototgtotg cetttoccet tagagaaget gtomttaga tgacctgato ccotgtoctg gcccaggetg gacaagtgoa gosostgosg otatttosag gttgcccoaa atgccctago ttggtattca taggggaggt tetgeetgee gtgccaocco gtagocggag asatcccatg scaggogtgt gtgcccaggt accetgeee tggosasota aggagggtas gocaggacco tottgoctgg accattgggc otgotttggg tytoctysto actggtggcc tagagggoca ctaococact cactettqtt agctgggatt cttcaggtga acaatttaag gtetestgee ggtggagota ctctgcctgg ogstyttotg gototgetee cotottttct tacctgtoto tosasasas saccocasas sasasacco aggeoottgt gttgcaotgg ctoctocast occtocceto gggetttgge ecoagoetee ttggtagotg.tgaggogtca ccotccagat geographtt cagggactta ttggcccttg cacattotog ggatgaagcg cttctgocco gggcagccg cotocoaagt ttagcoagsa gatttotcaa tacagtgaga geagectete gatggttgtt ctgtggette asscreeds tgagggcagg tgtctccago tossgagagt gatottotat gggetgetgg aggeetgtgg atggatcoat astecaggee gtecetaste cotgcotoag gcctgggaca ttocggaaga aocagggcag tttttgtgga ggetggtete tettggesce gtgggggtcc gtttcatatg ttactgattg assaggtass ctattaatct tgaaggttgo agggetgtg gttettetag teceagettt actttaaget cctttactca tettggttet tagcagggga ccacgggggc goccea tgco coagtigigo tgttttttt aagogattot gaagaggaac agagtocagg agacgtgagt ggageteagt ttctccagge ttgtetetta ectgeetget atgttggtca toctattttt tgctageact gagggcctco cctgtggccc gggagactea tgeseteett cettgttett tggagtgcat ggcctgccta cagtgacago gaggadttoc actoactgas tetetetetg tecegggtte aggittetee gtgocaggoc aagaggtgo cggaattaga tagocctgct ogtttettgt atattcatga tgggtgtacg ttggtcotca ggtoagagee totgttgtct tctcagtcaa agtacctgoc actgragtgt accttccag cagtegtage ggtggggga gggaaggett gtoaggageo ascetotgeo stotocotoc stgootgtga cagotttggg ptggaagaot gesesetggs cotgettoca agtagagaca ptgagccaco gagtaaaggg aaagagggt ttggcactgg cgoagccago cctgggattt caetacaaag cacagaggag gaacgaggag cogsaagsaa caggtaggtg acgrasocot gggoaatggg ctggcottgt coagggaag **80800088800** -3429 -3329 .3229 3129 3029 2929 -2829 2629 2729 -2529 2429 2329 2229 -2129 -2029 -1929-1829 -1729 -1629 -1529 -1429

EXON IB	EXON IA	.·	EXON II	EXON III	EXON IV	6B
coccyctoc agricotoc caeactgaec taagaaacce teageacaga caacaccect GCATTCCCCA EXON IB AGGAGGGCAC GAGGCCAAGAT TCCCAAGAGC TCAGGTGAGT gacacaccg aatggccag gacgccctca gaagcogagg cototgttat etotgccote tecccattga tateccatt cagacaacce cygcoggcet gaagcogagg cototgttat etotgccote tygagtgetg tyggatcetg gccaetgoag cettgaattc ggggaggca gytettgctg teacogagge tygagtgetg tagaaaatat ttttaaaaga tttttagaga getaggacta cagaccetea ceatectgc tygatagttt taaaaaatat ttttaaaaga tttttagaga	tocasattot ggootoagoo tocotagggt otgggattac ayytygysys TTTGTCCTGG AACCCAGATG EXON IA cocacatoag gototggoo ocaataccag otoagttgot caatgggotf TTTGTCCTGG AACCCAGATG EXON IA cocacatoag gototggoo ocaataccag otoagttgot caatgggotf TTGTCAGGAGG GGAACATCGG GGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA	CCTGGGATTT ctgactggcc taagggaacg	TCAAGTACTT K Y F googcocacc			FIG. 6B
cottoottoo ggaageaaat otaagteeag ecceggetee agateeetee eaestgaee taagaaacee teageacaga caacacceet [echinicacka <u>CACAACACCC ACACTCABCC ACTGCGGGCG AGGAGGGCAC GAGGCCAGGGT TCCCAAGAGC TCAGG</u> Ggag gacgeccag gacgecotes <u>CACAACACCC ACACTCAGCC ACTGCGGGCG AGGAGGGCAC GAGGCCAGGGT TCCCAAGAGC TCAGG</u> GGGG ecctgeteag ettgtggete eaaeatteea gaageegag oetetgttat etetgeete toccatgga tateccatt cagacaace eggeette gaateeceet coetteett tittititee ggggaggeg gyettigetg teaeogagge tggagtgetg tgggateetg gecaetgeag ettttagaga gaateeceet coetteett tittititiee ggggaggeca ggtettgetg cacacagage tggatagtt taaaaaatat tittaaaaga titttagaga	tocasattot ggootoagoo tocotagggt ctgggattgot oaatgggotb TTTGTCCTGG AACCCAGATG cocacatoag gototggooc ccaataccag ctcagttgot caatgggotb TTTGTCCTGG AACCCAGATG cocacatoag gototggocc ccaataccag ctcagttgot caatgggotb TTTGTCCTGG GGAAGAGGAGGAGGAGGAGGCCAGCCAAGCCA	GAAATGGTGA GACAAAGGGC GCIGGCIGGC GCTTTCACTG GTCTGCACCA TGGCCCAGGC CCTGGGATTT CCTGACATCT GGGCCCCCGTT CCACGAGGGT GCTTTCACTG GTCTGCACCA TGGCCCAGGC CCTGGGATTTT GGGCGGGGCTG GGGAACCACO ACALLAGAAC CCGAGCCGGTC LLCCAGCCTCL GTGALGGTGT LAAGGGGGG	ctaggectes graetures graetures carresagna ceresaases erccectra reastarit resaacresa casesacresa casesacresa crassacrit X Y F E I O D I S R P P L E Y V K G V P L I K Y F F E L I O D I S R P CASCINACCC AMERICASTA ggragg geoscosse	CCSGGCACCA T T	TGACCTGGAA AAGTGTCACC GAGCTCCCAT CITCATGCG GTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCCTGAAGA CACACCTGCC CCTGGCTCTG CTCCCCAGA TCTGTTGGA TCAGAAGGTC AAGGGGGGGTCTG CTCCCCAGA TCTGTTGGA TCAGAAGGTC AAGGGGGGGTCTG T L D Q K V K  L L K T H L P L A L L P Q T L L D Q K V K  F G.
teageacaga gacacacegg tateceattt tgggateetg taaaaaatat	osstoggete CTCAAGGAGG	GTCTGCACCA tttcagococ tcttqctgtt	CGTGAAGGGG	ocotocotot	T P F L agacetteee	CTCCTGAAGA CACACCTGCC CCTGGCTCTG CTCCTGTTGGA TCAGAAGGT L L N T H L P L A L L P Q T L L D Q K F L L K T H L P L A L L P Q T L L D Q K F
taagaaacco TCABytgagt tocccatgga tggagtgctg tggagtgctg	otogostrac otosottoct GecceAGGG	GCTTCACTO cogsoctggt	CACTGGAGTA  L E Y  CACCTACCC	Q A R P D D L L I S T Y P K S SI Agogaggica tgotoacotc agoctgotoa cotocoatot ecotocot	rrcarscoe F M R toageteace	CTCCCCCAGA
cacactgaco TCCCAABABC ototgocoto tcacogaggo	tocotagggt ccastacoag ACATGTGAGG	CCACGAGGGT acattagaac	TCTCGCCCGC S R P P TGCTCATCAG	L I S	GABCTCCCAT A P I A ggctggggct	ccreecrcre
agatocotoo GAGGCCAGGT pototgttat ggtottgotg cagacoctoa	gootoagoo gototggooo AGGAGTTGCC	GGCCCCGTT GGGCCCCGTT GGGSSCCCC	CCAGGACATC  O D I  CCTGATGACC	P D D L	AAGTGTCACC	CACACCTGC
coccyyctoc AGGAEGGCAC gaagrogagg ggggaggoca qotaggacta				Q A R		CTCCTGAAGE
otaaqtcoaq ACTGCGGGCG Caacattcoa tttttttco	cagattggto tttgotcatt arcacaGecc	GTCCCAGCAG CTTCACACAC TGAAAGGTG	otoaacctoo cgcag5AACA	1 Cacttggoo	Accassesses 0 6 6 cctgggtgca	A P R
cottoottoo ggaageaat otaagteesg CACAACACC ACACTCAGCC ACTGCGGGCG coctgetesg ettgtggete caacattees gaateccoot coettoottt tttttttee	tagggtotto caatgotgoo osgattggto ttttgotgag toatcacagt tttgotcatt	CTACTGGGTG GAGGCTGGGG GTCCCAGCAG AGCTGGACTC TCCTGGAGAC CTTCACACAC TGAACAGCTC CGCAGGTGAA TGAAAAGGTGA	tigigagigo ggyoaagica otoaacetoo ggeetgyete iggeceeiga egeag <mark>SAACA</mark> H	TGCAGAGGCA CTGGGGCCCC TCCACAGGCC  A E A L G P L Q S F  ctctcccagg tggcagtccc caccttggcc	ccacarrer Gacargarer accassesce of I I I I I I Q G G B	I P S G GAGACACACC AGCCCCACGA GAGACTCTGA AAAACACAC AGCCCCACGA
cottcottco CACAACACCC coctgotcag gaatoccoct	tggggtotta tggggtotta ttttgotgag	CTACTGGGTG AGCTGGACTC	ttgtgagtgo ggootggoto	TGCAGAGGCA A E A ctotocoagg	CCAGATTCTG Q I L	I P S G GAGACTCTGA E T L K
-1229 -1129 -1029 -929	-823 -729 -629	1329	-129	27 271	272	472

:	·	EXON V	EXON VI	
tccaeaaaaa ttacagcgag ggctcacac ctctactaa ggcagaggtt gcaggtggca	cagtotogoco ggaaggagaa caaaacaaat gagttooaga tcaaaagtot gaaagcocaa ccaggtgtgg tcatottoca		ACTACCACTT CIACLANAIS OCCUPANTY OF H P G T W E S F L E K F  Y H F Y H H A K W Y Y P H P G T W E S F L E K F  gaaggaaggt gtggagetaa ggggtggtgg otacaacgca cagcaacct gtgtcggcac ccctgcccg  gaaggaaggt gtggagetaa ggggtggtgg otacaacgca cagcaacct gtgtcggcac  AGCACGTGCA AGAGTGGTGG GAGCTGAGC GCACCCACCC TGTTCTCTAC CTCTTCTATG AAGACATGAA  B Y Q E W W E L S R T H P Y L Y L F Y E D H K  H Y Q E W W E L S R T H P Y L Y L F Y E D H K  AGCACGTGCA AGAGTGGTGAG GCACCCACCC AGGCCACCC CCTGTGGGGGGCCCCC CCTGGGGGGGGGG	78 71
	testacoago actitograg giogagging goagatcaco caaggitagg agtitogagai cagicinggoc tacaaaaat agocaggogi ingggoaggo gocagiaato coagotacio gggaggoing ggaaggagaa inggcogaga iigitocaci giaciocago oigggagaca caacgagaca iigiticaaa caaaaaaa gotcacogoi giaaigocag caitiiggga ggocaagaig ggingacogo iigagcicag gagiticaaga totataaaag aigaagcaa cagoaggo aiggingoa agocaaaa gacgocaa caggaaaaaa gaaagcocaa toaacgiigo agigagotai toteacicca gigoacicca acoigggoaa caggaaaaaa gaaagcocaa gacciagagi occocococ aaaaaaaa aaaccacaa aaaaagaaaa aagcaaaggi ocaggigigg gacciagagi occocococ aaaaaaaaa aaaccacaa caaaagaaaa aagcaaaggi ocaggigigg	otcagoccag circygroot groups octaagaag gragagytg gytcotcaca gaccacaaa gaccaaaaa atgagactot tocggytaga cotaagaag gyattatatt actatagaat atgatgytet agcatgoott gytcotacto gygytytgac groottaaga atotcaaaca agtcagocto taagcotcag caggyttcag atcetgactt gyccotacot gygocyttga gtottotgcc cetgcaggrG GTCTATGTTG tygootgygg caggotgtgt gytgatgytg otgygyttga gtottotgcc cetgcaggrG GTCTATGTTG tygootgygg caggotgtgt gytgatgytg otgygyttga gtottotgcc GAAAGCTTC TGGAGAAGTT	ACTACCACTI CIACARATO COMPANY Y P H P G T W E S F L E K F Y H F Y H F Y P P P T W E S F L E K F Y T W P W T W T T T T T W T T W T T W T L F Y E D H K W W T T W T L F Y E D H K W W T T W T L F Y E D H K W W T T W T L F Y E D H K W W T W T L F Y E D H K W W T W T W T W T T W T W T W T W T W	
magttocag a tgagoctag g lagtgtoca o caacatggo a gaattgott g	cagottagg a cagotacto g aacgagaca t gtggaccgo t gcgctatag t cotgggcaa c	ragaggytg ggattatatt a totosaaca a fottotgcc o	G T W E E E FEFFETCTAC C T T T T T T T T T T T T T T T T	
gtaotttggg aggetgaggt gggaagatoc ettgaageea gaagtteeag ataagtetet gtaataeeag ttaeteagga ggttgaggtg ggaggateat otgageetag gagttaaagg gtgaeagaga aacactgtot caaaaaacga tgaatagaaa gagtgteeea eeagtgeggt aggtggatea eetgagaeta ggagtttgag ateageetgg ecaacatgge aaaaceeea aggeatetgt aateeeaget acttgggagg etgaageagg agaattgett gaagetggga	tastacoago actticogneg giogagotogo gocaquaco caaggitago agiticoga tacasasast agocagoogi ingogoago gocaquasto ccagotacto gogagotoga togagosqua caacogogoca tigiticoasa goticacogot giastoccag cattitogoa gocaaquig gotigacogo tigiticoasa totatasasag atgosastas teagosagoca atgotogoat agogotatag teccagotac teacogito agitigagotat tetoscicca gigosocicca acetogogoca cagosasasa tosacotaco assassasas asaccasco toscicco cagotatoco cagotico	octaegaagg goattgatgo a gtgoogtgag a otggggttga g	PHP otaceacges of SCACCCACCC T THP TSSOOTGES	
gggaagatoc c ggttgaggtg c caaaaaaga t ggagtttgag a acttgggagg c	togaggtog tygggcaggo gtactocago catttggga tcagcaggo tctoactoca tctoactoca	tooggtaga ggcctact ggtgatgtgc ggtgatggtg	A K Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y	
aggetgaggt ; ttactoagga ; tacactgtot ; cotgagacta ; aateccaget ;	tataccago actticogne gicongotogo tacaaaaatt agocagoot togogocago togogocago tactoccago gotocago totataaaag atgoaaataa toagoaggo toaacogtoo agigagotat totaactoca quectagagic tocaacogocago atgoaaataa totaacagogo toaacogoca	atgagactot gotcoagato atcotgactt caggotgtgt	T H H Graysagetaa AGAGTGGTGG W W Colored	
gtaotttggg aggetgaggt gtaataceag ttaetoagga gtgaeagaga aaoaotgtot aggtggatoa eetgagaeta aggestetgt aateceaget	tastacoago tasaaaaatt tgagccgaga gotcacogot tctataaaag tcaacgttgo	otcagoccag ctruggicory gaccasasae atgagactet agcatgoott gotccagate cagggitcag atcctgactt tggootgggg caggetgigt	ACTACCACITY Y H F gaaggaaggt gaaggaaggt AGCACGTGCA H T Q	i i i i i i i i i i i i i i i i i i i
	·			تقودودودود
igttcacacc jtgcatagtg icoagtgcac jcacttgaag iattagccgg	otgitigged of the standard of	coagggaag gtctgctt ggtggaaca gcagactt	CGCAACGC AAAGGATGTG GCGGTTTCCT  R N A K D V A V S Y  TGGCTGGA GAAGGTGGG ttgatgggag  I A G E Y  Lotccagf GTCCTATGGG TCCTGGTACC	cogocttoga
tyggcacagt ggttcacacc cgcaatotca aaacttagct gtgcatagtg gtgtgtgcot ctatgatcac accagtgcac tccaggctgg tgtaattcca gcacttgaag aggctgaggc aatacaaaa aattagccgg gcatggtggc	gaggactaga cotattata ctecocota sacatagasa ascocotot otectasas tagettacae etagrasgog gtggttgcag saatattta asaggtttgc oscotogytg cospeceage sacataggg agactecate saggttggag agactecate cagettggag gattgcttga goctgggagg agattettt etetttete tttttttga	gostgigs tecsggssg gaggeeeegg eggggggse tgietgeett tigeteigst sittigiggi ggiggsses igeeiggigg eagggseesg gesggseit ggeittigsg etteeteet igeessees sgsgsigsg	CCCGCAACGC AAAGGATGTG GCGGTTTCCT R N A K D Y A Y S Y CATGGCTGGA GAAGGTGGGG ttgatgggag H A G E N H A G E N CATGCCAGT GTCCTAGGAGC cttctccagT GTCCTATGGG TCCTGGTACC	GGAGgtgaga cogocttga tgctccccc
572 672 772 872 6972 6			2372 0 2472 0 2572	2672

FIG. 6C

```
1 acctctgcct cctggttcca agcaatcctc cttcctcacc ctccagagta gctgggatta
 61 cacgcgcctg ccaccgcgcc tggcctaatt tttgtatttt tagtagagat gggggtttcc
121 aaccatgttg gccaggctgg tctccaaact cctgacctca ggtgatcctg cccacctaag
181 cctcccaaaa tgctggtatt acaggcatga gccaccgtgc ccggcctaaa taattaataa
241 aataatggac gatgggtgcc ttctactgag ctcccggtaa ttgtgagtga gtagaggact
301 tgccctgggg acattcagtg acctgctggg tgttgctgag ctgtgaggaa gttcaggtct
361 ggctgcagtg gtgaggctgt gactcaatca atcactgctg atgctcccag gacctgcacc
421 agettagtee taggggcaag gattttaact gtccacetca gtctcttcat ttgtaagatg
481 caaataacag tcacccctgc ctcatgggat ggagctgtgt aatgcccgca acagtgcctg
541 ctgcatagag gggttgctgc cagctgcctc tecetecttg tetettacet geetgetgcc
601 tagatcada taaaqaacaa coottgtgtt gcccccacc tagctgcctg ctaagggccc
661 atgtgatctg cctggcagag gagtttcttc aggaagaacc agggcagctt ctgcccctag
721 agggccaatg cccttggtga gtgcagtccc ctggccccag cctggtccac ctctgggaag
781 agggtgccca gttgtgcaat ccaggcccag gcagctgage cctcatctca gcatgcaggg
841 cggatactgg agggggettg tggcatetga etetgtatet cetacetgce ceteteettg
901 gradetgra gaagteactg ettrggggag acetgatetg getgreeng arggacactg
961 agaaagaagt agaagactca gaattagaag aggtgagtgg gctttggtgg cgggctccct
1021 accocacted otgeoctygy otgeotytya coacactyct tycototyca gycacactyg
1081 acagacetge tggagacetg atceteagtg teettaceee etectacete tittetgtge
1141 cacctgctgt gggtccagca ggtttttact tgagtacaat aaaaagtctg agtcaagggt
1201 gccttatggt ggatgctgag gggaggggcg gagctagtag cccaaggtcc tgccagtcac
1261 ggggcttcct caggggcaca gaggaggcag gaggggcccc tggccctagc acgtgaacag
1321 cttctactct gcctggaaac cccatgcctc agctttcccc tacttgcctc tgagctcatg
1381 caattett99 aageet999a gaettaeett gaaattgaat geaaatagga caaagaeeaa
1441 ggaggatggg gggatgccct ccttccacgg ggccctgtgg cttccaagtc ttaatctcct
1501 ctagtctctt gtctacggag cctccttcaa acccagggaa agaaaagcac ctgccagggt
1561 tgtttttctt ctaggatett ctattgatge tetgtgaggt cececaggag ccatgaaget
1621 agggctggct cctagggcaa tgggactaca gtgtccttgt cctttcttat tctttctgtt
1681 ctttcttttt ttttttttt tttttttt ttttttgag acagagtete actetgttge
1741 ccaggctgga gtgcagtggt gtgatcttgg ctcactgaaa cctccgcctc ctgggttcaa
1801 gtgattctct tgcctcagcc tcctgagtag ctaggattac aggtgcccgc catcatgccc
1861 agctaatttt tgtattttta gtagagacag ggtttcacca tgttggccag cttggtctcg
1921 aacteetgae eteaggtgat eetgetgeat egaceteeca aagtaetggg attacaggeg
1981 tgagccacca egeteageet etttettgtt etatatgtee atgetetget ecaettetge
2041 cccttcactc tgccccacac atcactccag actggccttg tggtcagagc ctggaatgcc
2101 tgggctgctg ggggcctgtg gactgcactg ggccagaacc cctgccgcct tcaagactgg
2161 cctgtagcca gcaggtaggt gacttttccc aggccggcct atcccacctt tcccctccac
2221 teacteacet ecettgeetg ggteaattag agaaagettg teggeeagge atggtggete
2281 atgcctgtaa teteageact ttgggaggee gaggegggeg gateatetga geteaggagt
2341 ttgagaccag cctggccaac atggcaaaac cccgtctcta ctaaaaatac aaaaattaac
2401 cggatgtggt ggtgtgcacc tgtaatccca gctactcggg aggctgaggc agaagaatcg
2461 cttgaaccca ggaggggag gttacagtga gcggagatcg tgctactgca ttgcagcctg
2521 ggcgagagag cgagtctcca tctcacataa aaaaaagaaa aagaaagaaa gcaagcttgt
2581 ctgttggcct gccctgcagg gtggagttca gagggaaggt caggagccta gtgacagctc
2641 aaaaaaaaa aaacccaaat accaatgttg gccccttttg cctttcattc atgtgttttc
2701 tatacactaa actcacatat tgggtttgca gatcactcca agcttggctg gagctgtggt
2761 ggtaaggagg gtaatagaga agetteecea eeeteaacee caccettee tteetggagt
2821 teccageet gaetttagat eceteegaea etggaeette aaaaceetea gggeagagag
2881 Cagcoctaca otocotacao cacacceata otoageceet geaggeaagg agagaacagg
2941 teaggtteec gagageteag gtgagtgaca egetggaatg geecagggea cetteaceet
3001 getcagettg tggetceaac attetagaag cegaggeete tgecatecet gecettteee
3061 atggatatte catticaatt agacaaceca geetggeegg aacececetg egiteettet
3121 tttcctttgt gtatttttga gacagggtgt tgctccgtca cccaggctgg agtgtagtgg
3181 gatectggcc cactgcagec teaaatteet aggetgagge aatectgeeg ceteageete
```

```
3241 ctgagtagct ggggttacaa gagcaagcca ccacacccag ctaattttga aaaatattt
3301 ttgtagagga gaggtettge tttgttgtee aggttggtet caaactecag ggeteaaggg
3361 atcettecce gttggeetee caaggetetg ggattacagg egggagteac eetgeetggy
3421 coccecttt tgatgagtca teagttttea tteeegeacg aggetetage ecetggtace
3481 agettagttg ctcaatggge tgtgtttgtt etggageeea gatggaetgt ggeeaggeaa
3541 gtggatcaca gacctggccg gcctgggagg tttccacatg tgaggggcat gagggggct
1601 caaggaggg agcatcgggg agaggagcgc actgggtgga ggctgggggt cccagcagga
3661 aatggtgaga caaagggege tggetggeag ggagaeagea caggeaggee ctagagette
3721 ctcagcacag ctggactctc ctggagacct tcacacaccc tgatatctgg gccccgcgct
3781 acgagggtgc tttcactggt ctgcactatg ccccaggccc tgggattttg aacagctctg
3841 caggtgactg aaaggtgcgg ccaggctggg gaacgacctg gtttcagecc cageccegec
3901 actgactgac titgtgagtg cgggcaagtc actcagcctc cctaggcctc agtgacttcc
3961 ctgaaagcaa aaactctgca aaggggcagc tgggtgctgg ctcacacctg taatcccagc
4021 actttgggag getgaggtag acaaateact tgaggeeagg agreetagae cageetggee
1081 aacatggtga aaccccatct ctactaaaga aaaaaaaaa ttagctgagc atggttgtac
4141 atgettgtaa teccagetae ttgggatgee gaggegggag gattgettga acccaagagg
4201 tggagtttgc agtgagctga gattgtgcca cactgcactc cagcttgggt gagagtgaga
 4261 ctccatctca aaaaaaaaaa aaaaaagaga gaatcccact ttcttgctgt tgtgatggtg
 4321 graagggaac gggccrggct crggccccrg argcacgaac arggagcrga rccaggacac
 4381 ctcccgcccg ccactggagt acgtgaaggg ggtcccgctc atcaagtact ttgcagaggc
 4441 actggggccc ctgcagagct tccaagcccg acctgatgac ctgctcatca acacctaccc
 4501 caagtctogt aagtgaggag ggccacccac ceteteccag geggeagtee ecacettgjt
 4561 cagcaaggte gtgeeeteag eetgeteace teetatetee eteetetee acgeaceace
4621 tgggtgagee agataetgga catgatetae cagggeggeg acetagagaa gtgtaacegg
 4681 geteceatet aegtaegggt gecetteett gaggteaatg atecagggga acceteaggl
 4741 gcatggctgg gtcctggggg taagggaagt ggaggaagac agggctgggg cttcagctca
 4801 ccagacette ectgacecae tacteagge tggagaetet gaaagacaea cegececeae 4861 ggeteateaa gteacaeetg eccetggete tgeteeetea gaetetgttg gateagaagg
 4921 teaacatgag geegggetea atggtteaca cetgteatee cagtttgaga etgaggaggg
 4981 aggatecett gaaggegaga gatggagace ageetgggea acattgetgt agagatgaea
 5041 teccatetet acaaaaataa aattaacaac etggtatggt ggcatagact gtteccagtt
 5101 acttaggagg ctcagcgggg aggactgttt atgcaaatag gaagctgcaa tgagccctga
  5161 tgatectget getgeactee ageetgggea acacageaaa accateteta egaaaaaaaa
  5221 agttcccact gactggcaag gaaagccagg aaggggggct caggtgccct ctcagccatg
  5281 tacctgttct totggaaggg cotcotoget totgcoagge toatcacate tttttttt
  5341 ttgagacaga gtcttgctct gtcaccctgg ctggagtgca gtggcatgat ctcagctcac
  5401 tgcaacctcc gcctccccag ttcaagtgat tctcctgcct cagcctcctg agtagctggg
  5461 attacagers telectacca caccegeta attettetat telettagt agagacegeg
  5521 tttcaccatg ttggtcaagt ggatctcaaa ctcttgacct tgtgatcctc ctgcctcgac
  5581 ctcacaaagt gctggaatta caggcgtgag ccaccgcgcc tggccctttt ttttttgag
  5641 acagtttcac tottgttgcc gaggetagag cgcaatcgtg tgatctcggt teactgcaac
  5701 caccgcctcc tgggttcaag caattctcct gcttcagcct cccaaggagc tgggattaca
  5761 ggtacctgcc-accacgcccg gctaattttg tatttttagt agagatgggg tttcaccatg
  5821 ttggtcaggc tggtcttgaa ctcctgacct caggtgatct ggcaccttgg cctcccaaag
  5881 tgccgggatt agaggcatga gccaccacgc ccagcettca tcacatettg agagaggaca
  5941 ctgtctgcct cttgctctga tgagggtctg atgcaaagga tagtgagtct ctacagtgca
  6001 cacttaagaa aggcagcatg tgggtgctca caggtcaggc ggaggagggg gagctggtgg
  6061 ggaccaggca tgccttgctc cagatcagga tatgatggca ttggtgcaga ttatattagt
  6121 atagaatatg gtctcaggaa ccaggcagga ctttggcttc cgagcagggt tcagatccca
  6181 gettggeet acctgtgeag tgagatetea ageaagteag cetetaagee teaggtteet
   6241 cetttgecag tteaacagat gagetggeet ggggtggget gtgtggtgat ggtgetgggg
   6301 ctqqqtcctc tqcccctqca qqtggtctat gttgcccgaa acccaaagga cgtggcggtc
   6361 tectactace atttecaceg tatggaaaag gegeaceetg ageetgggae etgggacage
   6421 ttcctggaaa agttcatggc tggagaaggt gggcttgact ggaggaagga gggtgtgaag
   6481 ccgaggggtg gtggctataa cgtacagcaa ccctgtgtcg gtgccccctg cccgcttctc
   6541 tadegeceta egggecetgg taccageacg egcaggagetg gegggagetg agecgcacec
   6601 accordated chacetette tatgaagafa tgaaggaggt gagacegact gtgatgette
```

```
6661 cccccatgtg acacctgggg gcaggcacct cacagggacc caccaaggcc acccagccc
6721 gtccctgggc ggctcccaca gcaagcccgg attccccatc ctacctccct ggcccaggcc
6781 ccccactgc agcccacct ggcagcaggc tcggcacagc tttcatcttc tgcacctgag
6841 tcagctgcat gggtggccac ggatcagata cttagtccta ttgcttatcc tcaccaaag:
6901 gtgtgccacc cagggccaca gtcatggaag aagaccatcc cggtcctcac ccataggcgc
6961 caagccctgt tcatgatggg atcacagggc agagatcaat tcattttact cca_kgacta
7021 gggccccagg ggttgaggct ctttggggtt tctaggggaa gtggccagat cccctctgag
7081 gttagagagg gggacccgtt ttgttttgct ccactgagga gccctctgct gctcagaacc
7141 ccaaaaggga gattcaaaag atcctggagt ttgtggggcg ctccctgcca gaggagacca
7201 tggacttcat ggttcagcac acgtcgttca aggagatgaa gaagaaccct atgaccaact
7261 acaccaccqt cccccaqqaq ctcatqqacc acaqcatctc ccccttcatq aggaaacqtg
7321 ggtgctggcc agcacggggg tttgggggcgg gtgggagcag cagctgcagc ctccccatag
7381 geactt9999 cetecect99 gatgagacte cagetttget ecetgeette etececeach
7441 catggctggg gactggaaga ccaccttcac cgtggcgcag aatgagcgct tcgatgcgga
     ctatgc99a9 aagatggcag gctgcagect cagetteege tetgagetgt gagagggget
7561 cctggagtca ctgcagaggg agtgtgcgaa tctaccctga ccaatgggct caagaataaa
7621 gtatgatttt tgagtcaggc acagtggctc atgtctgcaa tcccagcgat ttgggaggtt
7681 gagetggtag gateacaata ggecaegaat ttgagaecag eetggtaaaa tagtgagaee
7741 tcatctctac aaagatgtaa aaaaattage cacatgtget ggcacttace tgtagtccca
7801 gctactt999 aagcagagge tggaggatca tttcagecea ggaggttgtg gatacagtga
7861 gttatgacat geceatteae tacageetgg atgacaagea agaceeteee tecaaagaaa
7921 ataaagetea attaaaataa aatatgattt gtgtteatgt agageetgta ttggaaagga
7981 agagaaacte tgagetgaaa gagtgaatge ceggtgggge cacatatggt cacetetece
8041 ccagcettca getececagg teaceatate tggggagggg agaagggttt ggagaagtaa
8101 aacccaggag atgtgtggag gggggatgte tgtttaatce cagcacatce tetgetgtee
8161 tgccccaaga tggtggagga cgtcgagtcc gccgggcagc gtcacttttt cttgggctcc
8221 ttagaagcta ccaggtacct ctgggccaca ctgagatgag gggagtagcc gcctgcatag
8281 gaggtgtett caaacaggat agtatagtee eteetggggg ttgtgggggt aggtggecaa
8341 ggaagggtag aggagcaagc ccccggggct ggttgtcaac tcactttgtt ggctggaatt
8401 ggttgtaact tgaccacctc gggcaggatc ccactgctca tccccaa
```

FIG. 7C